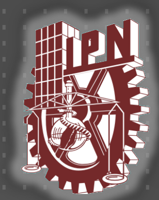
Instituto Politécnico Nacional

Centro de Investigación en Computación



Retos actuales del procesamiento de lenguaje natural en la inteligencia artificial moderna



Dr. Hiram Calvo

Coordinador de la subcomunidad de procesamiento de lenguaje natural

¿QUÉ ES INTELIGENCIA ARTIFICIAL?

Razonamiento

Sistemas que *piensan* como humanos	Sistemas que *piensan* racionalmente
Sistemas que **actúan** como humanos	Sistemas que **actúan** racionalmente

Comportamiento

¿PUEDE PENSAR UNA MÁQUINA?

- Prueba de Turing:
  - Una máquina será inteligente…
  - …cuando sus respuestas a las preguntas de un operador humano…
  - …no sean distinguibles de las que daría una persona.

- IA como imitación del comportamiento humano…

PROGRAMAS DE DIÁLOGO – ELIZA, 1966

```
==========================================================================
          EEEEEEEE  L         IIIIIII  ZZZZZZZ      AAA
          E         L            I          Z      A   A
          E         L            I         Z      A     A
          EEEEE     L            I        Z       A     A
          E         L            I       Z        AAAAAAA
          E         L            I      Z         A     A
          EEEEEEEE  LLLLLLLL  IIIIIII  ZZZZZZ     A     A

==========================================================================
ELIZA > (PLEASE TYPE IN ALL CAPS) WHAT'S YOUR NAME DEAR ?
AMIT
ELIZA > HI, I'M ELIZA. WHAT DO YOU WANT TO TALK ABOUT ?
AMIT > CAN I TALK ABOUT MY PROBLEM ?
ELIZA > SURE... !
AMIT > DUE TO ON GOING EXAMS IN COLLEGE, I HAVE A LOT OF STRESS.
ELIZA > PLEASE RELAX A LITTLE. SLEEP WELL.
AMIT > THANKS FOR YOUR ADVICE.
ELIZA > NO MENTION.
AMIT > BYE
ELIZA > BYE AND KEEP IN TOUCH...

==========================================================================
```

CHATBOTS

A.L.I.C.E. and judge

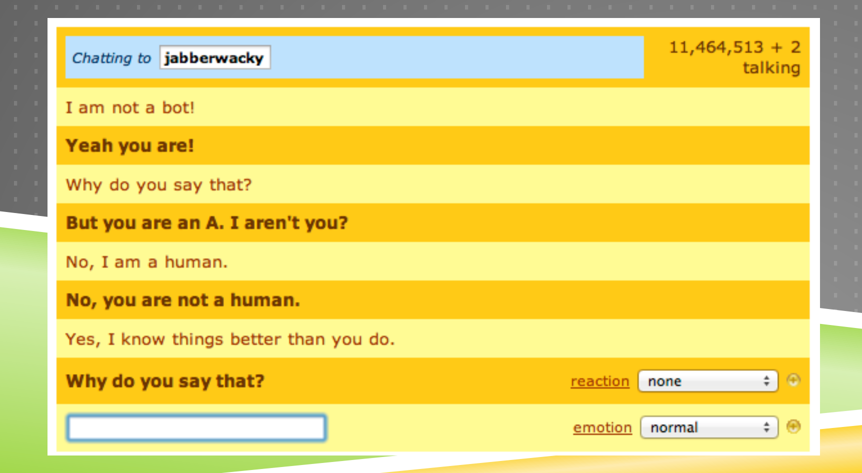
You said: You answer very quickly.
A.L.I.C.E.: My responses are determined by my personality.

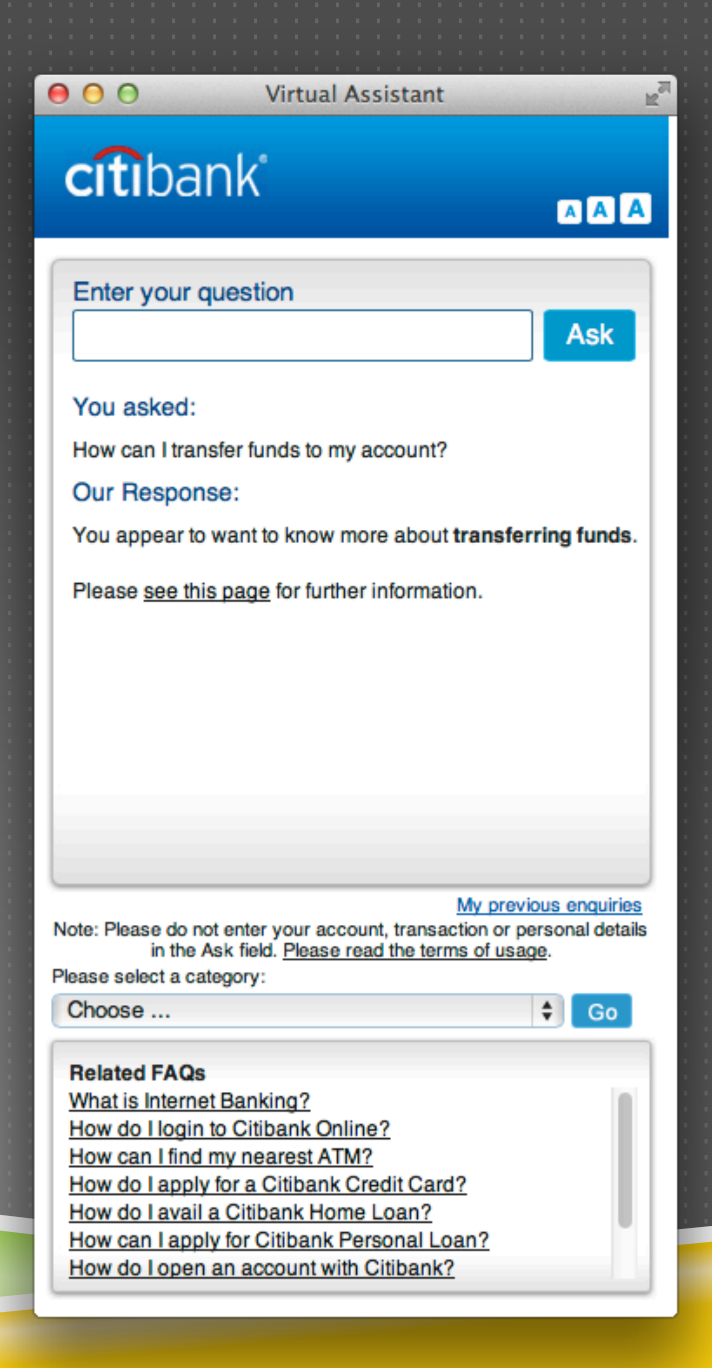
You say: Say

Learn how this works.

Conversation Log:

judge: Where are you from?
ALICE: I am originally from Bethlehem, Pennsylvania. Now I live in Oakland, California. Where are you?
judge: What are you doing?
ALICE: I am just killing time.
judge: What is your favourite hobby?
ALICE: Chat robots make an excellent hobby.
judge: Are you a chat bot?
ALICE: What made you think I might be a robot?
judge: You answer very quickly...
ALICE: My responses are determined by my personality.





ASISTENTES DIGITALES

- Un conocido asistente se comunica en lenguaje natural con los usuarios…



¿QUÉ SE LE PUEDE DECIR?

- ¿Qué opinas de la inteligencia artificial?
  Déjame que lo piense.
- ¿Se encontrará la partícula de Dios?
  Mi política me obliga a separar el espíritu del silicio.
- ¿Existe Dios?
  Todo esto es un misterio para mí.
- ¿Cuantos años tienes?
  Tengo la edad suficiente para ser asistente.

¿QUÉ SE LE PUEDE DECIR?

- ¿Qué opinas de los partidos políticos?
  Pienso, luego existo, pero no descarto lo contrario.
- Eres una tonta
  ¿Después de todo lo que he hecho por ti?
- ¿Estás enamorada?
  ¿Tú crees?
- ¿Eres humana?
  En la nube nadie sabe si eres un ser humano o no…

- *"Existirá Inteligencia Artificial cuando no seamos capaces de distinguir entre un ser humano y un programa de ordenador en una conversación a ciegas"*

Alan Turing, 1950

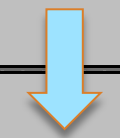


El cuarto chino (John Searle, 1980)

INTELIGENCIA ARTIFICIAL

Razonamiento

Sistemas que *piensan* como humanos	Sistemas que *piensan* racionalmente ?
Sistemas que **actúan** como humanos	Sistemas que **actúan** racionalmente



Comportamiento



"Artificial intelligence is overrated. We find most people prefer the user-friendly advantage of artificial stupidity."

TRADUCCIÓN AUTOMÁTICA (1960s)

inglés → ruso (o español)→ inglés

"The spirit is willing, but the flesh is weak" →

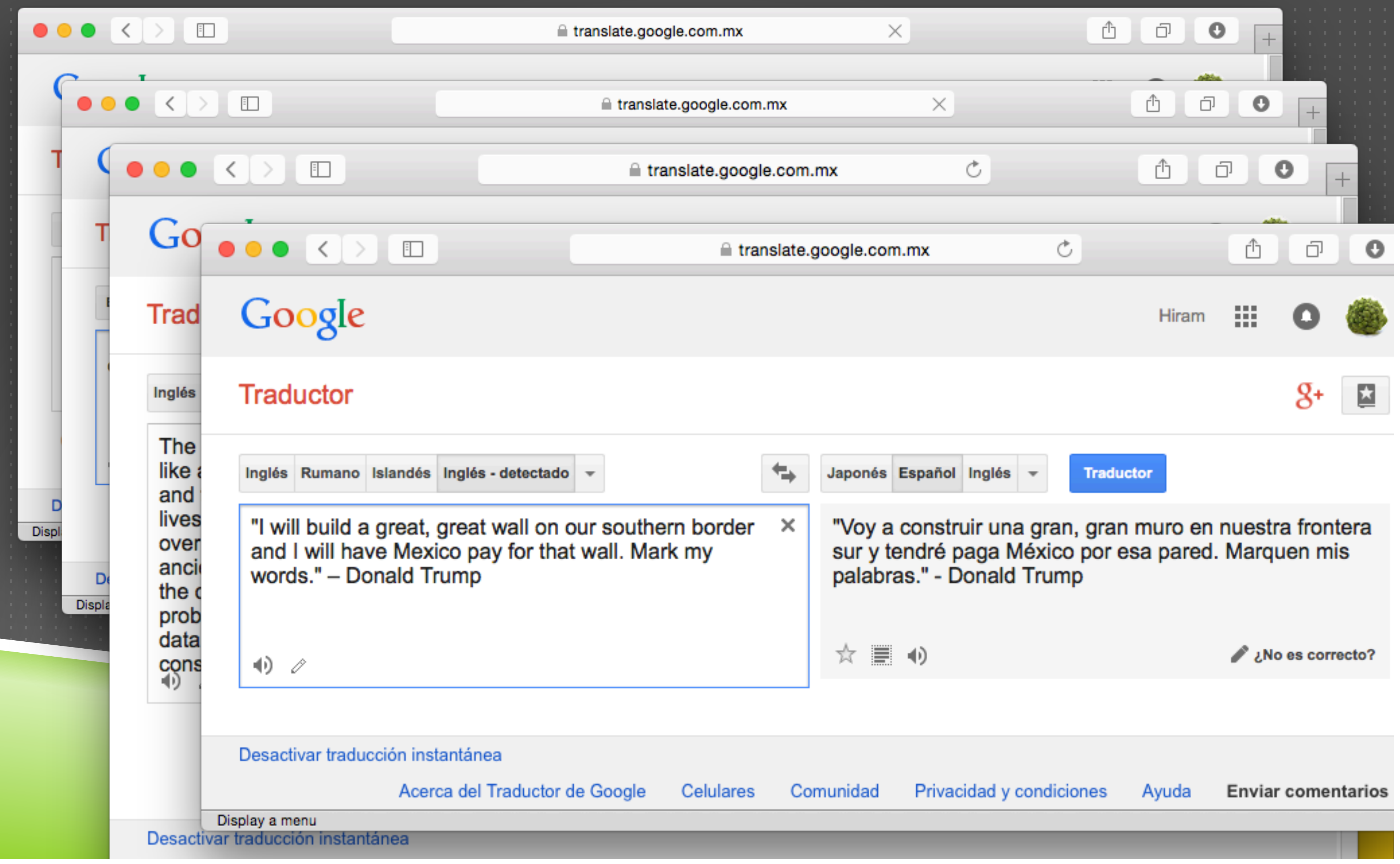
→ Дух охотно готов, но плоть слаба →

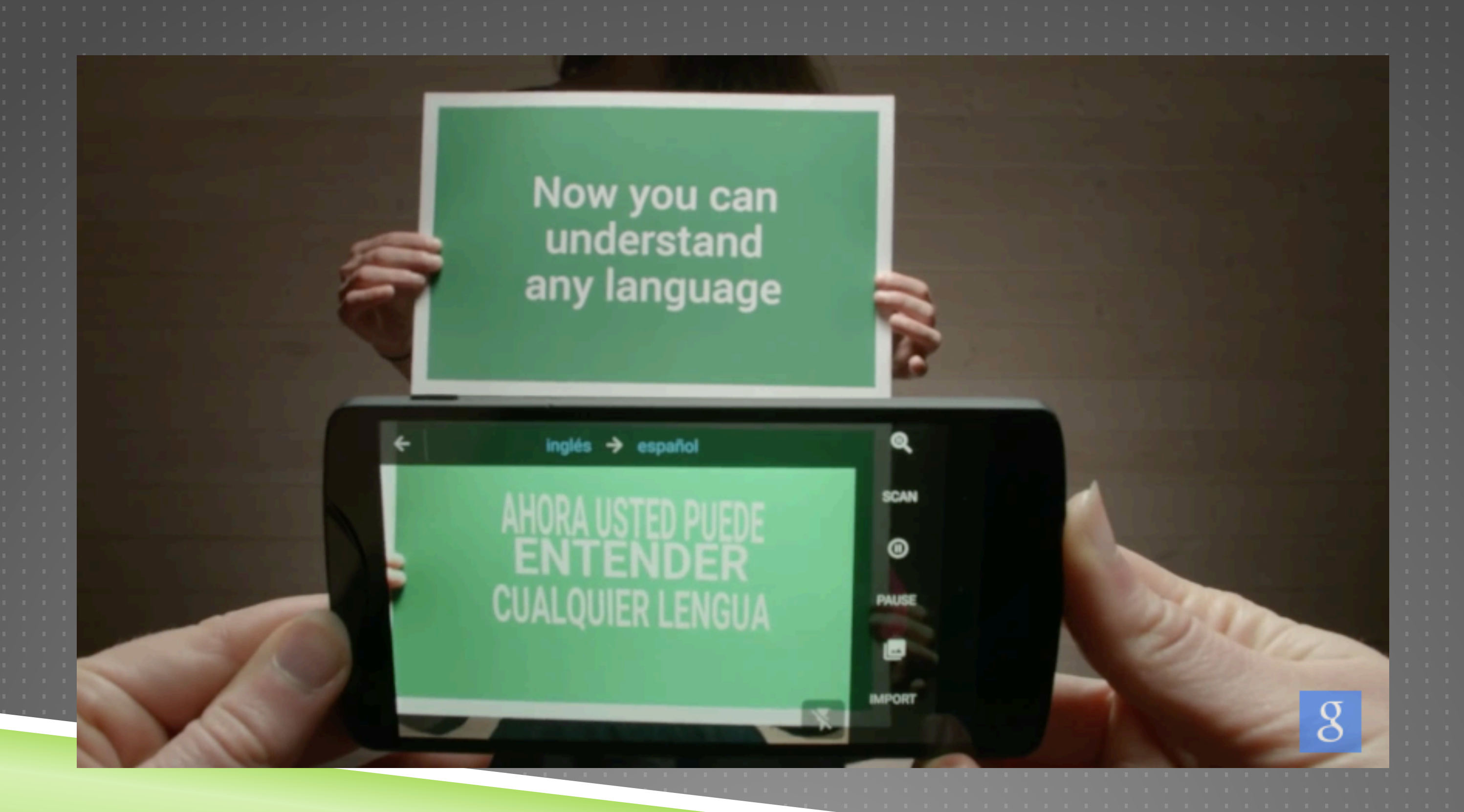
→ "The vodka is good, but the meat is rotten"

(Alternativas:
"Spirit is willingly ready, but flesh is weak"
"The alcohol is arranged, but the meat is weak")

les enfants et les femmes enceintes
→ pregnant children and women

HOY EN DÍA (EJEMPLO: TRADUCCIÓN DE GOOGLE)




Now you can understand any language
inglés → español
AHORA USTED PUEDE ENTENDER CUALQUIER LENGUA
SCAN
PAUSE
IMPORT

mail.google.com

what is the gratest problem in artificial int

google transla... | Funny Google... | Expo 2015 'G...

You Should Neve Google Translate

By Nick Cowen on 27 May 2015 at 3:45AM

For some, the existence of Google Translate – tools – means that they think they'll never aga language. Well, guess again.

While online translation apps are good for get a foreign language, it turns out that they're rat hold a conversation with anyone.

In his latest video YouTuber Tom Scott takes u that occur with online translation and points o certainly capable of matching words and con work with, it misses out a lot of the subtle nua favourite example in the video is when Scott English person saying "that's a brave idea!" c

Why Computers Suck At Translation

gizmodo.co.uk

google transla... | Funny Google... | Expo 2015 'G... | translate error... | Translate ya p

DEBISTE NUNCA CONFIAR EN TRADUCTOR DE GOOGLE (Here's Why)

By Nick Cowen on 27 May 2015 at 3:45AM

For some, the INMINENCIA GOOGLE Translate – Y UN CURTIDO OTRO ORDINE INMOLACIÓN CONDENAR MEDIO that they think they'll NUNCA again TIENE QUE HOGAR PENDIENTE DE NUEVO language. Well, ADIVINAR GANAR.

While EN LÍNEA PERSONA ENTERADA BOAS SE BUENO PODRIDO NO UN SENTIDO LO QUE ESTÁ CERVEZA MADERA FUERA in UN foreign EQUIPAJE, it turns out that they're rather useless if USTED WADI A USO them A CÓMO UN CONCURSO ion with ALGUIEN.

EN HA latest video YouTuber Tom ESCOCÉS takes NOS through EL ALGUNOS DE EL problems that occur with ANIME translation Y HACER in FUERA QUE while Google Translation ES certainly capable of CORRESPONDIENTE words Y PEINAR ARRUINADO it PROHIBIR EL CONTEXTOS TA DESPERTADO h, it misses out a lot of the subtle MATICES CONVENTO MAMÁ LENGUA NUESTRO favourite example in EL VÍDEO COMEDOR ESCANEAR PE COSTILLA out QUE UN American Y UN INGLÉS person saying "that UN OLA GUAU DAN't MBAR the same thing.

Why Computers SuCEBO ranslation

"TODOS NUESTROS MODELOS ESTÁN MAL"

- Pero la pregunta práctica es qué tan mal tienen que estar para no ser útiles
  - George E. P. Box y Normal R. Draper, 1974





RETOS ACTUALES

- Manejo de expresiones idiomáticas:
  - En holandés: at spille kong gulerod significa: ¿quién se cree que es?
  - Traducción literal: jugando rey zanahoria

  - John kicked the bucket: Juan estiró la pata.
  - Traducción literal: Juan pateó la cubeta.

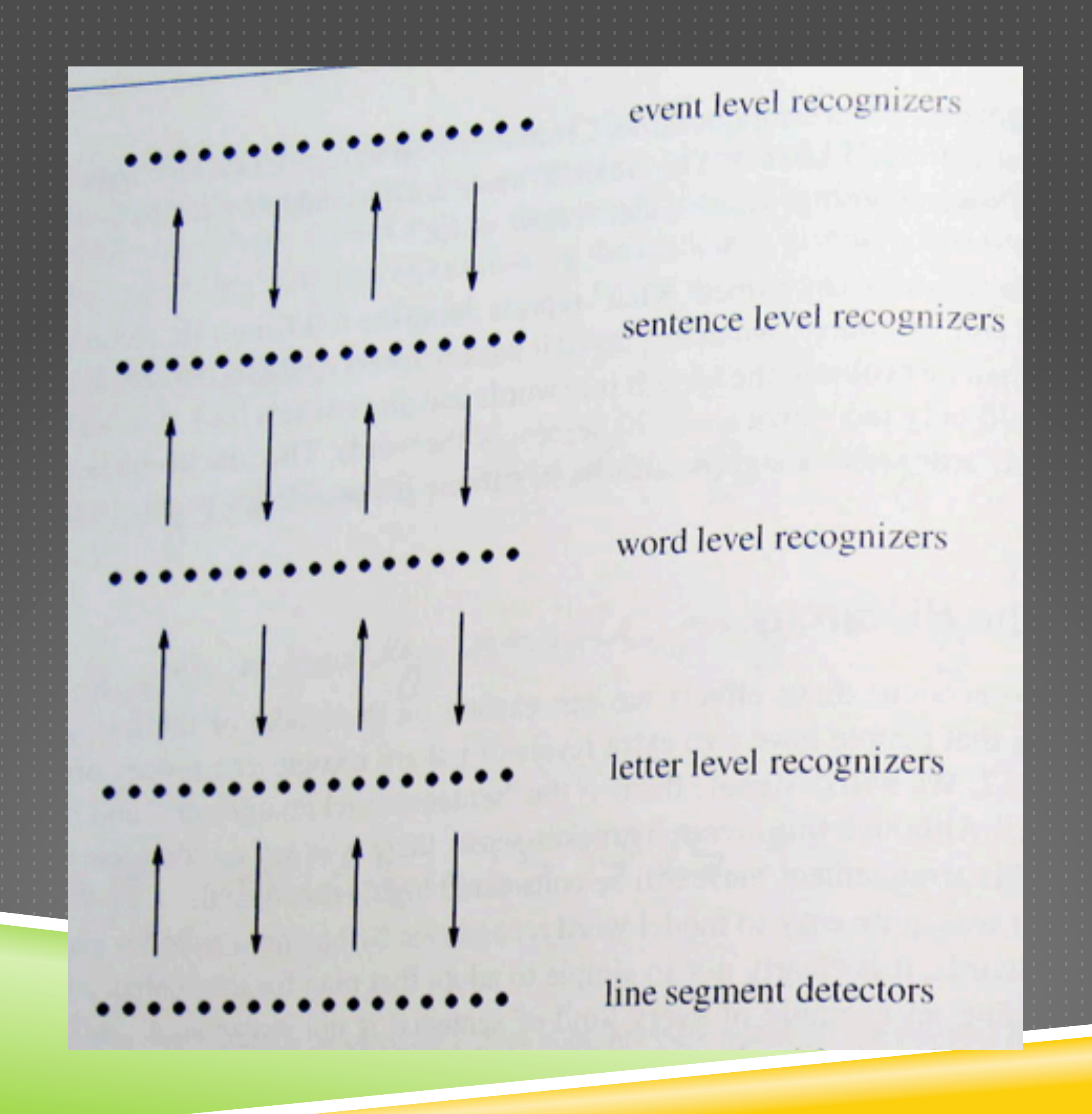
- Ambigüedad

2.4.2 Higher Levels

His feet are made of clay.

That really made my day!

THE PRICE IS $15.95.


event level recognizers
sentence level recognizers
word level recognizers
letter level recognizers
line segment detectors

PROBLEMAS QUE SE CONSIDERAN *CASI* RESUELTOS

- Etiquetado de categorías gramaticales (98%)
  - JUAN COME EN CASA
  - SUSTANTIVO VERBO PREPOSICIÓN SUSTANTIVO

Part-of-Speech Tagging from 97% to 100%: Is It Time for Some Linguistics?

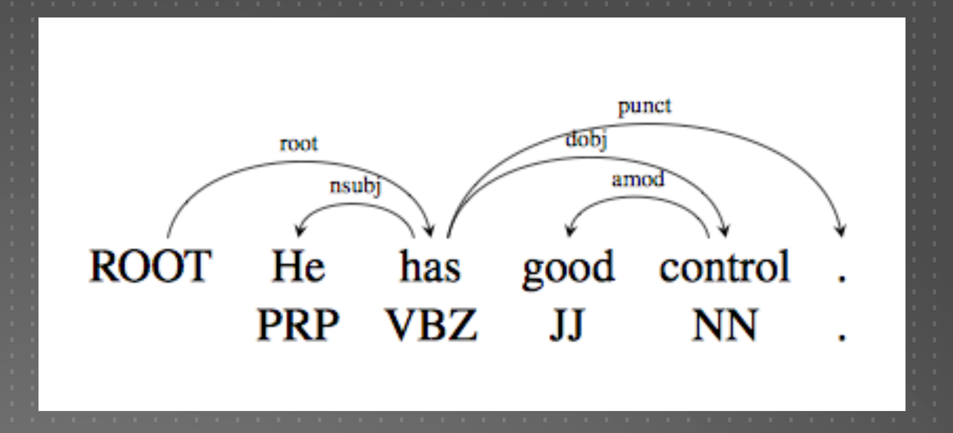
Christopher D. Manning

Departments of Linguistics and Computer Science
Stanford University
353 Serra Mall, Stanford CA 94305-9010
manning@stanford.edu

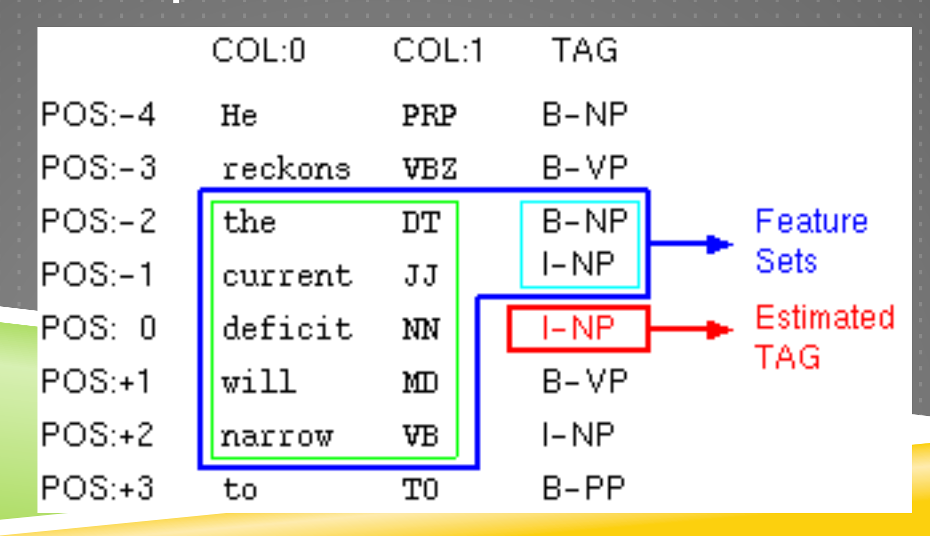
Abstract. I examine what would be necessary to move part-of-speech tagging performance from its current level of about 97.3% token accuracy (56% sentence accuracy) to close to 100% accuracy. I suggest that it must still be possible to greatly increase tagging performance and examine some useful improvements that have recently been made to the Stanford Part-of-Speech Tagger. However, an error analysis of some of the remaining errors suggests that there is limited further mileage to be

ANÁLISIS SINTÁCTICO (EN INGLÉS)

- Ha superado el 92%



- Análisis sintáctico superficial, 95+ %



	COL:0	COL:1	TAG
POS:-4	He	PRP	B-NP
POS:-3	reckons	VBZ	B-VP
POS:-2	the	DT	B-NP
POS:-1	current	JJ	I-NP
POS: 0	deficit	NN	I-NP
POS:+1	will	MD	B-VP
POS:+2	narrow	VB	I-NP
POS:+3	to	TO	B-PP

RETOS ACTUALES QUE ATRAEN EL INTERÉS INTERNACIONAL

- Paráfrasis y semejanza semántica en textos cortos
- Semejanza semántica textual
- Representación semántica de un texto
  - Identificación de eventos
  - Identificación de orden de eventos en un texto
  - Identificación de causas y efectos
  - Identificación de espacios y lugares
  - Resolución de correferencia
  - Extracción de grafos conceptuales

RETOS ACTUALES QUE ATRAEN EL INTERÉS INTERNACIONAL

- Análisis de opinión masiva
  - Considerando textos con lenguaje figurado – metáforas
- Desambiguación de los sentidos de todas las palabras en una oración
- Análisis de taxonomías (ontologías)

¿QUÉ HACE NUESTRA COMUNIDAD DE PROCESAMIENTO DE LENGUAJE NATURAL?

- Detección de textos sospechosos (¿son reales las opiniones vertidas en algún sitio?
- Agrupamiento de textos noticiosos de múltiples fuentes
- Predicción de reacciones en redes sociales a noticias
- Medición de distancia semántica a través de dispositivos BCI (ondas EEG)
- Sistemas de consulta legales

- Interfaces de interacción humano-computadora



MILENIO
Diario

Detección de plagio

Expertos del IPN ganan primer lugar con "software" antiplagio

El modelo desarrollado, superó a competidores de otras naciones, como Chile, Estados Unidos, España, Alemania, China y Reino Unido.







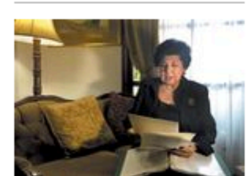




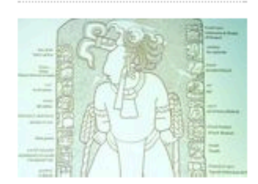






Porfirio Díaz transformó a México, afirman sus bisnietos



Descifran nombre de tumba del rey Pakal gracias a glifo maya



De Porfirio Díaz, la primera imagen filmada en México





Miguel Ángel Sánchez Pérez es estudiante de doctorado en el IPN. (Especial)

REDACCIÓN
05/01/2015 02:14 AM

México • El alumno de doctorado, Miguel Ángel Sánchez Pérez y los investigadores Alexander Gelbukh y Grigori Sidorov, del Centro de Investigación en Computación (CIC) del Instituto Politécnico Nacional (IPN), desarrollaron un modelo de detección de plagio que permite identificar textos producto de la piratería.

Con ese *software*, los politécnicos obtuvieron el primer lugar en la categoría de Alineación de Textos de la 11 edición del *Evaluation Lab on Uncovering Plagiarism, Authorship, and Social Software Misuse* (conocido como PAN) celebrado en la Universidad de Sheffield, Inglaterra.

RESPUESTA AUTOMÁTICA A EXÁMENES

Table 2. Overall results for all runs, QA perspective

RUN NAME	C@1	# of questions ANSWERED				# of questions UNANSWERED
		RIGHT	WRONG	TOTAL	Prec.	
Synapse-French	0.59	33	23	56	0.59	0
Synapse-English	0.45	25	31	56	0.45	0
DIPF-7	0.38	21	35	56	0.38	0
cicnlp-8	0.38	21	35	56	0.38	0
cicnlp-7	0.36	20	36	56	0.36	0
csgs-1	0.36	20	36	56	0.36	0

- 2o. lugar competencia internacional CLEF

- Respuesta automática de series numéricas en tests psicométricos

(respuestas dentro del promedio)

RECIENTEMENTE…



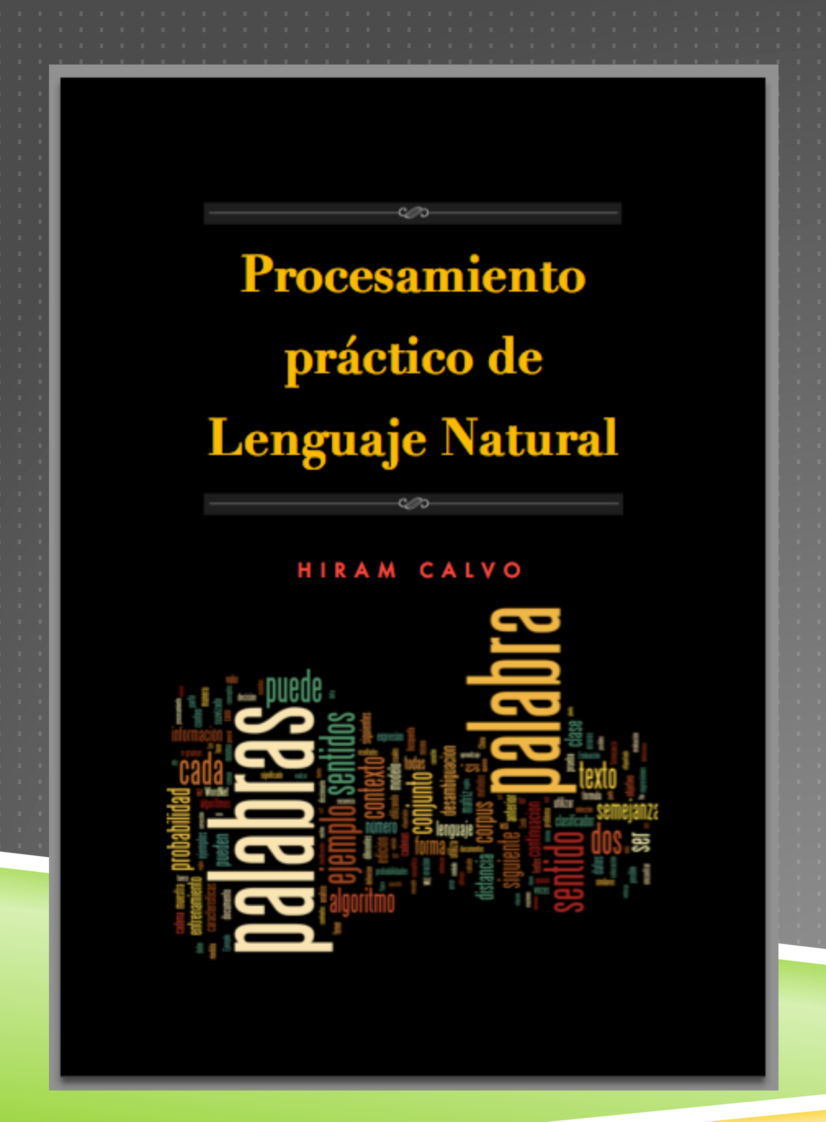


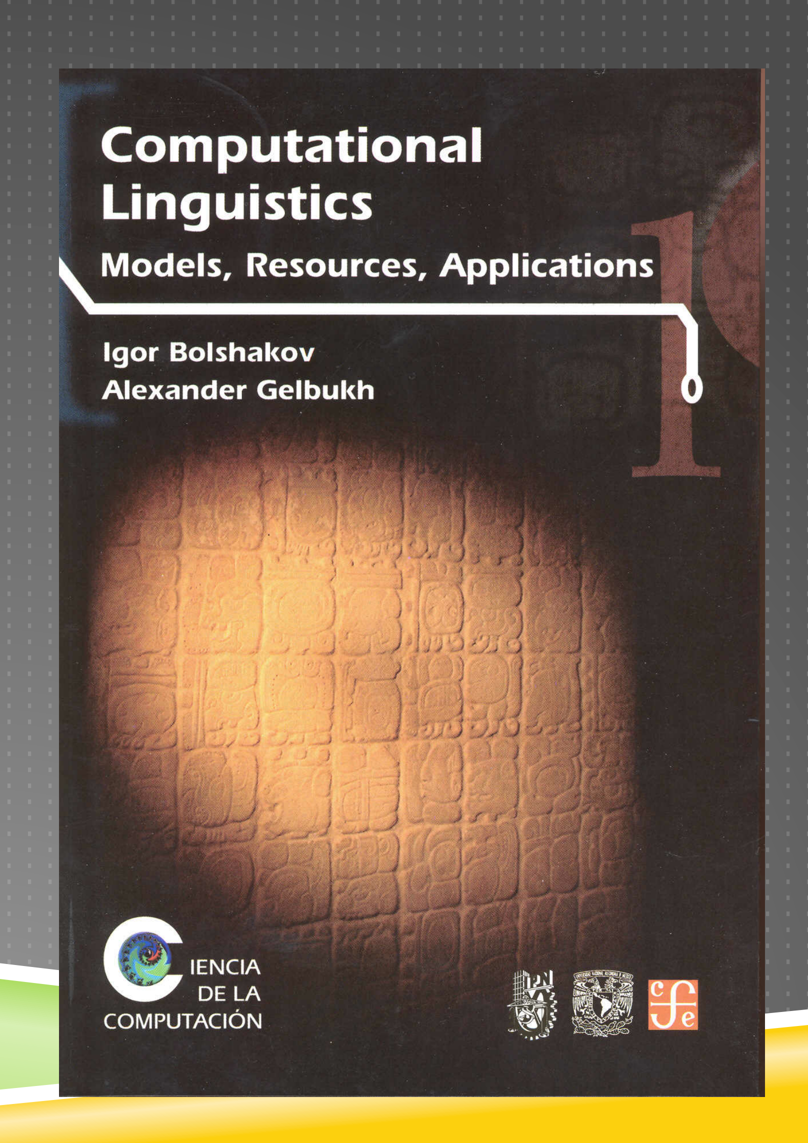


"Resultados de la Convocatoria para la continuidad de Redes Temáticas CONACYT 2015"

- Fue aprobada la red temática en Tecnologías del Lenguaje, por Conacyt
- Pertenecemos a la subcomunidad de procesamiento de lenguaje natural de CUDI

MÁS INFORMACIÓN…





¿QUÉ SIGUE?

- Mark Johnson de la universidad de Macquarie:

Prediction is very difficult, especially about the future

– Niels Bohr

- My main prediction for the future:
  Computational linguistics will be so successful that in 50 years . . .
  it may not exist as a discipline any more
- I hope I'm wrong


Gracias…

RESPUESTA A SERIES DE TESTS PSICOMÉTRICOS

serie	Individuo								Computadora		Resultados
	1		2		3		4				
1	562	n(i)= n(i-1)*3	567	*3	567	-	567	*3	567	3*A	567
2	30	n(i)= n(i-1)+1	30	1	30	-	30	+1	30	1+A	30
3	162	n(i)= n(i-1)	162	*3	162	-	162	*3	162	3*A	162
4	4096	n(i)=- 2^(i-1)-5	-1024	a/2	1024	-	-2048	*4	-2048	.4*A	-2048
5	12	-	13	-	12	+5	12	-5	12	-10+A	12
6	32	-	32	-	32	(4,1)	32	-1	32	3+A	32
7	8	-	8	-	8	(1,2)	8	-3	8	A-3	8
8	24	-	24	(-5,-2)	24	(5,2)	24	+3	24	3+A	24
9	9	-	9	(-2,-4)	9	(-2,-4)	9	-6	9	A-6	9
10	44	-	44	(5,3)	44	(5,3)	44	+8	44	8+A	44
Total aciertos	18		22		20		25		21		

RAZONAMIENTO FORMAL

- "Mecanización del pensamiento".
- Una vieja búsqueda:
  - Métodos deductivos en las antiguas China, India y Grecia.
  - Los silogismos de Aristóteles.
  - Los "Elementos" de Euclides.
  - Los algoritmos de al-Khwärizmi.
- Ramon Llull (1232-1315) desarrolla sus máquinas lógicas.
  - Máquinas para crear conocimiento combinando verdades innegables mediante operaciones lógicas simples.
- En el XVII, Leibniz, Hobbes y Descartes exploran la sistematización del pensamiento racional.
  - Hobbes: "La razón no es más que cálculos".
  - Leibniz: Reducir la argumentación a cálculos.
- Siglo XIX: Álgebras de Boole y Frege.
  - Formalización del razonamiento matemático.

RAMÓN LLULL (1232-1315)

- Llull es una de los primeros en tratar de hacer deducciones lógicas de una manera mecánica
- Su método fue un intento temprano de usar medios lógicos para producir conocimiento



ARS MAGNA (1315)

- Según Llull, la máquina podía probar por sí misma la verdad o falsedad de cualquier postulado que se le introdujese

